

A Secondary Screen Architecture to Accurately Capture Viewers' Interactions in an iTV Environment

Ricardo E. V. de S. Rosa
Graduate Program in Electrical Engineering
Federal University of Minas Gerais
Av. Antnio Carlos 6627, 31270-901,
Belo Horizonte, MG, Brazil
Email: ricardoerikson@ufmg.br

Lucas C. Cordeiro and Vicente F. de Lucena Junior
Graduate Program in Electrical Engineering
Federal University of Amazonas
Manaus, AM, Brazil
Email: {lucascordeiro,vicente}@ufam.edu.br

Abstract—Advances in TV technology have enabled viewers to actively interact with the TV via interactive applications instead of just passively watching TV. Typically, a conventional remote control is shared by many viewers and it is difficult to accurately capture the individual interaction data from each viewer. It is also difficult to identify the context in which the interactions occur, e.g., to manage precisely who is present in the environment and what is being watched on TV. In this paper, it is presented a novel architecture that facilitates the capture of viewers' individual interactions and contextual data in shared TV environments. This architecture uses personal devices as secondary screen devices with the aim of identifying viewers and capturing their interactions with interactive TV devices. A work in progress implementation of this architecture is briefly reported as an audio-visual content rating system for interactive TV.

I. INTRODUCTION

TV watching is essentially a social activity, where groups of people with a common interest share the same space and TV set for entertainment or information. After many advances in technologies, for interactive and digital TV, it is possible to develop high-level applications to enrich the watching experience. Thus, the interaction between viewers and TV became more elaborated than simply change channels and adjust sound volume. In this scenario, the data generated by the interactions of millions of viewers can provide valuable information for content providers (CPs) to improve their services [1].

The conventional remote control (RC) is typically seen as the standard medium of interaction among the viewers and their TVs. However, this medium presents two notable problems: (1) to precisely and automatically identify the viewers who have possession of the RC; and (2) to capture contextualized viewers' interactions. As a result, the entire group of viewers (e.g., a family) is treated as a single viewer and the tastes that were inferred from the users in charge of the RC can be imposed on the tastes of the others.

Given the limitations of the conventional RC, a novel secondary screen architecture to identify viewers who are interacting with interactive TV (iTV) services is presented in this paper. Personal devices (e.g., smartphones and tablets) are used as secondary screens with the aim of capturing personal and contextualized interactions from each active viewer in iTV environments. By using proper machine learning algorithms, the personal and contextualized data can reveal important information about viewers' tastes [2], [3].

Since personal devices are almost ubiquitous, they stand out as a powerful mechanism to identify viewers by means of wireless communication technologies [4]. In the architecture described in this paper, personal devices play an additional role as a medium of interaction since they are easy and intuitive to use in iTV environments, as pointed out in [5]. An approach described in [1] has shown an architecture to capture viewers' interactions via a conventional RC. The architecture described in this paper goes further and captures individual and contextualized data. Since the TV environment is shared, the data about nearby viewers is captured to contextualize the interaction with the purpose of helping to infer the behaviors of one viewer toward others ones.

II. SECONDARY SCREEN ARCHITECTURE

Initially, the CP infrastructure must provide an authentication mechanism so that viewers and their interactive TV (iTV) devices (e.g., set-top boxes or smart TVs) can be authenticated by the CP in order to use the secondary screen services. Once one iTV device is authenticated, the viewers in the TV environment must link their user accounts up with the iTV device. As a result, the viewer can use their personal device as a secondary screen, for that iTV device, in order to capture the viewers' interactions.

The architecture to precisely capture viewers interactions and the context is shown in Fig. 1. This architecture consists of three subsystems that interact with each other: (1) secondary screen devices, which might be implemented in personal devices; (2) an iTV device; and (3) the CP infrastructure, which includes an application hosting infrastructure. The dashed arrows represent the data flowing from the viewer to the CP infrastructure, while dotted arrows denote the flow from CP infrastructure to the viewers. The continuous arrows represent the data flow inside the CP infrastructure. Since applications in the TV domain have the potential to reach out millions of users at the same time, it is interesting to consider the use of a cloud based platform in the CP infrastructure [6].

A presence service detects physically nearby secondary screen devices that have the services necessary to establish a communication with an iTV device. Modern commercial operating systems for personal devices enable service discovery by using peer-to-peer or WiFi communication. This feature is particularly appropriated for the TV environment, once the

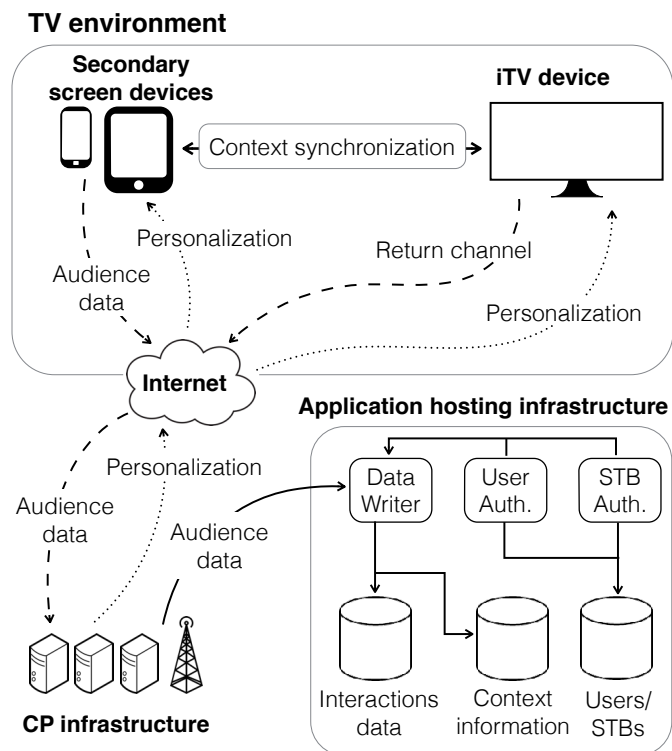


Fig. 1. Architecture to capture audience data. The viewers' interactions and the contextual information are captured and stored in the CP infrastructure.

secondary screen device must be wirelessly linked up with the iTV device. The devices are synchronized in order to update the contextual data about the viewer profile and the amount of people using a secondary device in the room. This contextual data can be used to learn how a viewer behaves towards other viewers. For example, a viewer behaves in a particular way while watching TV alone and the same viewer behaves in a different way while watching TV in group.

Viewers can use their secondary screen device to interact with TV by evaluating, commenting, sharing, or even recommending TV content to other viewers. The data related to those actions are naturally provided by the viewers, as part of the interaction, when using interactive applications (developed by the CP) installed in their secondary screen devices. The interactions and the contextual data are represented by the audience data, which is sent to CP infrastructure through the Internet. Then, a Data Writer service stores the audience data related to each particular viewer and the iTV device.

The Data Writer service in the application hosting infrastructure consists basically of web services that provide interoperability among secondary screen devices, iTV devices and the CP infrastructure. The main advantage of this approach is the independence of operating system and programming languages used in the personal devices.

Once the audience data is stored, the CP can use this data to derive valuable information about users' tastes. This information can be used to provide personalized content to viewers based on the current context. Thus, the personalized content can be delivered through the Internet to the secondary screen device or even to the TV, depending on the context.

III. CONCLUSIONS

The use of personal devices as secondary screens arises as a new trend in TV environments. This paper presents an architecture to accurately capture audience data by using secondary screen devices and iTV devices. This architecture takes advantage of the ubiquity and interactive capabilities of secondary screen devices to identify viewers and provide feedback for CPs through viewers' interactions and contextual information. The captured data relates to a single viewer in a given context. By using proper machine learning algorithms, this data can provide valuable information about viewers. Knowing this information can be very useful for CPs to improve their services, aiming to increase the audience and to enrich user experience. As an example, CPs can deliver personalized content for a single person as well as for groups of people by using content recommendation algorithms.

An implementation of the architecture that is described in this paper is under development in a commercial cloud computing platform, which plays the role of the application hosting infrastructure. A web application that consists of a content rating system was developed to store viewers' evaluations toward audio visual content; a secondary screen client is under development. Some RESTful web services were developed to provide interoperability among the secondary screen devices and the web application. Thus, the interactions that are captured during the use of the secondary screen application are sent to the CP by means of HTTP requests to RESTful web services. Although the secondary screen application is not fully finished yet, some specific features have been tested to capture viewers' interactions as they occur.

The architecture described in this paper is also technically feasible to implement in Digital TV (DTV) systems and can be particularly interesting for CPs in DTV standards, e.g., the Brazilian (ISDB-TB) and the Japanese (ISDB-T). These standards support the use of mobile devices as part of the system. Since the mobile devices support the implementation of the DTV middleware, a cooperative exhibition of the DTV application can be managed with the iTV device, thus enabling the capture of viewers' interactions and contextual data.

REFERENCES

- [1] C. A. C. Teixeira, E. L. Melo, R. G. Cattelan, and M. D. G. C. Pimentel, "Taking advantage of contextualized interactions while users watch TV," *Multimed. Tools Appl.*, vol. 50, no. 3, pp. 587–607, Feb. 2010.
- [2] H. W. Kim and M. Y. Yi, "MovieMine: personalized movie content search by utilizing user comments," *IEEE Transactions on Consumer Electronics*, vol. 58, no. 4, pp. 1416–1424, Nov. 2012.
- [3] H. Shin, M. Lee, and E. Kim, "Personalized digital TV content recommendation with integration of user behavior profiling and multimodal content rating," *IEEE Transactions on Consumer Electronics*, vol. 55, no. 3, pp. 1417–1423, Aug. 2009.
- [4] P. A. Cabarcos, R. S. Guerrero, F. A. Mendoza, D. Diaz-Sanchez, and A. Marin Lopez, "FamTV: An architecture for presence-aware personalized television," *IEEE Transactions on Consumer Electronics*, vol. 57, no. 1, pp. 6–13, Feb. 2011.
- [5] C. Courtois and E. D'heer, "Second screen applications and tablet users," in *Proceedings of the 10th European conference on Interactive tv and video - EuroITV '12*, vol. 11. New York, New York, USA: ACM Press, 2012, pp. 153–156.
- [6] S. Lee, D. Lee, and S. Lee, "Personalized DTV program recommendation system under a cloud computing environment," *IEEE Transactions on Consumer Electronics*, vol. 56, no. 2, pp. 1034–1042, May 2010.